

SULLIVAN & CROMWELL LLP

TELEPHONE: 1-212-558-4000  
FACSIMILE: 1-212-558-3588  
WWW.SULLCROM.COM

125 Broad Street  
New York, New York 10004-2498

LOS ANGELES • PALO ALTO • WASHINGTON, D.C.

BRUSSELS • FRANKFURT • LONDON • PARIS

BEIJING • HONG KONG • TOKYO

MELBOURNE • SYDNEY

Via ECF

September 26, 2019

The Honorable Robert W. Lehrburger,  
U.S. District Court for the Southern District of New York,  
500 Pearl Street,  
New York, New York 10007-1312.

Re: Chen-Oster v. Goldman, Sachs & Co., No. 10 Civ. 6950 (S.D.N.Y.)

Dear Judge Lehrburger:

Plaintiffs' demand that Defendants "must identify" all responsive documents sprinkled among the more than 37 million documents that did not hit on search terms ("Letter" at 2, ECF No. 857)—needles in the proverbial haystack—is plainly disproportionate to the needs of this case. Such a review would cost, conservatively, an additional [REDACTED] dollars, on top of the more than [REDACTED] that defendants already have spent during merits discovery. Because "[c]ourts cannot and do not expect that any party can meet a standard of perfection in ESI discovery, . . . a party requesting discovery may not be entitled, under the rules of proportionality, to every single relevant document." *Blackrock Allocation Target Shares v. Bank of N.Y. Mellon*, 2018 WL 2215510, at \*7 (S.D.N.Y. May 15, 2018) (Pitman, J.) (internal quotation marks omitted). This is a paradigmatic example of "the burden or expense of the proposed discovery outweigh[ing] its likely benefit." Fed. R. Civ. P. 26(b)(1). Plaintiffs have not identified any deficiencies in the massive volume of documents already produced, so no additional review is warranted.

**Defendants Have Conducted an Enormous, High-Quality Review**

In November 2018, Plaintiffs proposed "a search term route" for ESI discovery, and in January 2019, "Plaintiffs first proposed search terms." (Pls.' Mar. 11, 2019 Ltr., ECF No. 695.) The parties then engaged in "extensive back and forth to reach agreement on those terms" without success. (*Id.*) The Court resolved the dispute by authorizing Plaintiffs to "modify [Defendants'] terms," and Plaintiffs did so, "expanding 'within' delimiters" and adding whole terms to which Defendants had objected. (Mar. 19, 2019 Order ¶ 1, ECF No. 708.) The search terms chosen by Plaintiffs resulted in a "massive number (1.63 million) of documents captured by Goldman's recent ESI searches." (May 24, 2019 Order at 1, ECF No. 745.)

At considerable expense—more than [REDACTED]—Defendants reviewed more than 1.28 million documents,<sup>1</sup> identifying over 61,000 responsive documents and producing more than 500,000 pages. (Romero Decl. ¶¶ 4, 8; LaBrie Decl. ¶ 6–7.) Defendants also implemented extensive quality-control processes, including, *inter alia*, multiple levels of review and sampling of responsive and nonresponsive documents. (May 24, 2019 Order at 2; Ex. 3 (Defs.' Sept. 19, 2019 Ltr.) at 2.) After the review, Defendants' vendor used a representative sample of the search-term hit population to assess review quality, and determined that there were between about 20,000 and 75,000 responsive documents among the 1.28 million documents reviewed. The vendor

---

<sup>1</sup> The review excluded 335,942 irrelevant large Bloomberg files. (Ex. 1 (Romero Decl.) ¶ 4; Ex. 2 (LaBrie Decl.) ¶ 6 n.2.)

concluded that by identifying over 61,000 documents, Defendants had found all or nearly all of the responsive documents—more than 80% of the highest possible number of responsive documents that could exist in the review set. (Ex. 3 at 2; LaBrie Decl. ¶ 8; Letter App’x A at 2.)

### **Additional Review Would Be Disproportionate, Unduly Burdensome, and Duplicative**

At Plaintiffs’ request, Defendants also “review[ed] a sample of documents from the ‘null set,’ *i.e.* exported custodial documents that do not hit on the agreed-upon search terms,” and calculated specific Plaintiff-designed “metrics” based on that sampling review. (May 31, 2019 Joint Ltr. at 1, ECF No. 750.) Contrary to Plaintiffs’ selective quotation (Letter at 1), Defendants agreed only to “*produce*” the metrics Plaintiffs demanded, not that the metrics meaningfully assess whether yet more discovery is needed. (*Id.* (emphasis added).) In the “null set” sample of 384 documents, Defendants identified just three responsive documents—0.78%—all of which are substantively duplicative of produced documents. (App’x A.) If this analysis shows anything, it is that the additional review Plaintiffs seek would be disproportionate because it would target a tiny percentage of documents among 37 million, would not add to the discovery record in this case, and would cost a shocking [REDACTED] on top of the over [REDACTED] Defendants already spent.

Courts routinely deny as disproportionate requests to search for small percentages of potentially responsive documents among enormous collections. *See Blackrock*, 2018 WL 2215510, at \*7 (“[c]ourts cannot and do not expect that any party can meet a standard of perfection in ESI discovery”). In *In re Biomet*, for example, the defendant used keyword searches to narrow 19.5 million collected ESI documents to 3.9 million documents for review. 2013 WL 1729682, at \*1 (N.D. Ind. Apr. 18, 2013). Sampling of the “null set” in that case showed that 0.55% to 1.33% of the unreviewed documents would be responsive, *id.*, comparable to Plaintiffs’ 0.78% in this case. Just like Plaintiffs here, the *Biomet* plaintiffs sought additional ESI review by arguing that the defendant’s use of keyword searches missed a large volume of responsive documents. *Id.* at \*2. The court rejected plaintiffs’ argument, finding that plaintiffs’ proposed process “would cost [defendant] a million, or millions, of dollars,” and therefore the possibility of “unearth[ing] additional relevant documents” was outweighed by the “additional burden on, and additional expense to, [defendant].” *Id.* at \*3.

The exact same reasoning bars additional review here. Defendants spent more than [REDACTED] to review 1.28 million documents, and reviewing even a fraction of 37 million new documents would, unquestionably, cost millions more. Plaintiffs claim that there are “approximately 293,000 total” responsive documents somewhere among the 37 million documents that were not reviewed. (Letter at 2.) According to Defendants’ vendor, on average responsive documents appear in families of 3.72 documents (LaBrie Decl. ¶ 7), so that if Defendants could (impossibly) review just families containing responsive documents, this would be another 1.09 million documents, costing over [REDACTED] (*see id.* ¶ 5; Romero Decl. ¶ 8). Of course, because a review cannot be limited to responsive documents, the actual cost would be much greater.

The marginal benefit of additional review—on top of the enormous volume of documents already reviewed—does not come close to justifying this burden and expense. *See Royal Park Invs. SA/NV v. Deutsche Bank Nat’l Trust Co.*, 2018 WL 1088020, at \*8 (S.D.N.Y. Feb. 12, 2018) (Moses, J.) (“Proportionality focuses on the marginal utility of the discovery sought.”) (internal quotation marks omitted). Plaintiffs do not identify *any* document request not addressed or category of documents not produced and therefore have not carried their “burden of proving a specific discovery deficiency.” *The Sedona Principles, Third Edition*, 19 Sedona Conf. J. 1, 123–24 (2018). Moreover, assuming the “null set” sample is representative of documents that

did not hit on search terms—an assumption that must be adopted if the sample is used to draw any conclusions about that population—the fact that all three responsive sample documents are substantively duplicative of produced documents and peripheral to this dispute (*see* App’x & Exs. 4–6) suggests that the same is true of any other responsive documents in the “null set.” In other words, the marginal relevance of the documents Plaintiffs seek is effectively zero, an independent reason to deny Plaintiffs’ request. Fed. R. Civ. P. 26(b)(2)(C)(i) (courts “limit . . . discovery” where “the discovery sought is unreasonably cumulative or duplicative”). Although Defendants pointed out in the meet-and-confer process that Plaintiffs are seeking duplicative documents (*see* Ex. 3 at 4–5), Plaintiffs do not address the issue in their submission.

Lacking any valid basis to demand additional review, Plaintiffs rest their case on a purported “recall rate” (Letter at 1–2), which is supposedly the percentage of responsive documents Defendants *found* from among all responsive documents that were *collected*. Plaintiffs calculate this number based entirely on (i) finding three responsive documents (0.78%) in the 384-document “null set” sample, (ii) extrapolating that finding to 37 million documents and concluding that 293,000 documents were missed, so that (iii) the 61,000 documents that were produced amounts to 11.82% of all responsive documents. Plaintiffs calculate their own misleading recall rate incorrectly—it is 17.34% rather than 11.82%.<sup>2</sup> More fundamentally, the recall rate is meaningless with a document collection of this size. Plaintiffs achieve their recall rate because they demanded that Defendants *collect* a staggering number of documents—a total of 55.1 million from 62 custodians over 19 years—allowing Plaintiffs to apply their tiny percentage (0.78%) to a massive number (37 million). The ability to manipulate data in this way is one reason why Plaintiffs’ own authority states that “the [recall] number in its own right shall not be dispositive of whether or not a review is substantially complete.” *In re Broiler Chicken Antitrust Litig.*, 2018 WL 1146371, at \*6 (N.D. Ill. Jan. 3, 2018). Rather, courts consider “the duplicative or marginal nature[] of any responsive documents identified.” *Id.* Where, as here, additional production would be burdensome and duplicative, Plaintiffs are not entitled to additional discovery.

Finally, Plaintiffs’ demand that “the Court should require Goldman to use TAR” over its objections (Letter at 3) is contrary to law. As Plaintiffs’ authorities hold, the Federal Rules “do not give the requesting party, or the Court, the power . . . to force the responding party to use TAR.” *Hyles v. New York City*, 2016 WL 4077114, at \*2 (S.D.N.Y. Aug. 1, 2016 (Peck, J.)).<sup>3</sup> Moreover, preparing over 37 million documents for a TAR process would take at least 5–6 weeks and over 100 hours of work. (LaBrie Decl. ¶ 13.) This enormous volume of documents would need to be spread across multiple review sites—which is far out of the ordinary for even large document reviews—and will still require attorney review of approximately 3.4 million documents. (*Id.* ¶¶ 13, 15.) The total cost of such a review would be over [REDACTED]. (*See id.* ¶¶ 15–16; Romero Decl. ¶ 8.) Should the Court nonetheless require additional “production of documents that can be identified only through re-commenced processing, . . . review, and production, [Plaintiffs] will have to bear the expense.” *In re Biomet*, 2013 WL 1729682, at \*3.

<sup>2</sup> Plaintiffs understate “recall” by nearly 50% by ignoring 22,173 responsive family members of documents that hit on search terms. (Ex. 3 at 6 n.6; Letter App’x at 1.)

<sup>3</sup> *Rio Tinto PLC v. Vale S.A.*, 306 F.R.D. 125, 127 n.1 (S.D.N.Y. 2015) (Peck, J.) (“courts have refused” to order TAR) (collecting cases); *In re Viagra Prod. Liab. Litig.*, 2016 WL 7336411, at \*1 (N.D. Cal. Oct. 14, 2016); (“[N]o court has ordered a party to engage in TAR”); (Pls.’ Mar.18, 2019 Ltr. at 2, ECF No. 702 (“Goldman has the right to choose search terms over . . . TAR”)).

The Honorable Robert W. Lehrburger

-4-

Respectfully,

/s/ Andrew J. Peck

Andrew J. Peck

of DLA Piper LLP (US)

/s/ Amanda Flug Davidoff

Amanda Flug Davidoff

of Sullivan & Cromwell LLP

cc: GoldmanSachs-Attysonly@outtengolden.com (via e-mail)  
Carson Sullivan, Esq. (Paul Hastings LLP) (via e-mail)  
Ann-Elizabeth Ostrager (Sullivan & Cromwell LLP) (via e-mail)

## APPENDIX A

Responsive Null Set Sample Document	Previously Produced Document Hitting On Search-Terms
<b>GS0924397 (Ex. 4A):</b> This is a July 2013 e-mail chain between HCM personnel Linda Fox and Caroline Heller discussing talking points for a manager evaluation pilot program in the Securities Division.	<b>GS0560458 (Ex. 4B):</b> On June 27, 2019, Defendants produced substantively identical talking points for the same “Manager Evaluation pilot launch” in the Securities Division in an e-mail thread containing Ms. Fox and Ms. Heller.
<b>GS0924405 (Ex. 5A):</b> This is a May 2014 e-mail from Nancy Labiner in HCM titled “Changes to 2014 Performance Reviews.”	<b>GS0490052 (Ex. 5B):</b> On May 30, 2019, Defendants produced an April 2014 e-mail forwarding an e-mail by Nancy Labiner titled “Changes to the 2014 Performance Review Process – Presentation” containing an attachment with the same information.
<b>GS0924400 (Ex. 6A):</b> This is a 2015 slide titled “Performance Reviews 2015 Priorities.”	<b>GS0463300 (Ex. 6B):</b> On May 22, 2019, Defendants produced a nearly identical slide titled “Performance Reviews – Analytics: 2015 Focus Areas.” The parent e-mail to this document hit upon the search terms.